Generative models
Generative models

- Generative Moment Matching Networks
- Generative Adversarial Networks (GAN)
Generative Moment Matching Networks

• Black board
Generative Adversarial Networks (GAN)

• Original paper:
  – Generative Adversarial Nets

• Authors:
  – Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio (2014)

• Organization:
  – Université de Montréal

• URL:
Generative Adversarial Networks (GAN)
Generative Adversarial Networks (GAN)

- Bengio: This may hold the key to making computers a lot more intelligent.
Generative Adversarial Networks (GAN)

- Bengio: This may hold the key to making computers a lot more intelligent.

- LeCun: The most important breakthrough, in my opinion, is adversarial training (also called GAN). This is the most interesting idea in the last 10 years in ML, in my opinion.
Different Applications
DCGANs for LSUN Bedrooms

(Radford et al 2015)
Vector Space Arithmetic

• Similar to word embedding (DCGAN paper)

(Radford et al 2015)
PPGN for caption to image

• From natural language to pictures

Oranges on a table next to liquor bottle

(Nguyen et al 2016)
Adversarial Learning

Generative Adversarial Networks

$$\min_D \max_G V(D, G)$$

$$V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))]$$

Credit: Mark Chang
Training
Generative Adversarial Networks

\[ \min_G \max_D V(D, G) \]

Credit: Mark Chang
\[ V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))] \]
\[ V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_{\text{z}}(z)}[1 - \log D(G(z))] \]

\[ = \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_z p_{\text{z}}(z) \log(1 - D(G(z))) dz \]
\[ V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))] \]

\[ = \int_x p_{\text{data}}(x)\log(D(x))dx + \int_z p_z(z)\log(1 - D(G(z)))dz \]

\[ x = G(z) \Rightarrow z = G^{-1}(x) \Rightarrow dz = (G^{-1})'(x)dx \]
\[ V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))] \]

\[ = \int_x p_{\text{data}}(x) \log (D(x)) \, dx + \int_z p_z(z) \log (1 - D(G(z))) \, dz \]

\[ x = G(z) \Rightarrow z = G^{-1}(x) \Rightarrow dz = (G^{-1})'(x) \, dx \]

\[ \Rightarrow p_g(x) = p_z(G^{-1}(x))(G^{-1})'(x) \]
\[ V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))] \]

\[ = \int_x p_{data}(x) \log(D(x))\,dx + \int_z p_z(z) \log(1 - D(G(z)))\,dz \]

\[ x = G(z) \Rightarrow z = G^{-1}(x) \Rightarrow dz = (G^{-1})'(x)\,dx \]

\[ \Rightarrow p_g(x) = p_z(G^{-1}(x))(G^{-1})'(x) \]

\[ = \int_x p_{data}(x) \log(D(x))\,dx + \int_x p_z(G^{-1}(x)) \log(1 - D(x))(G^{-1})'(x)\,dx \]
$$V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[1 - \log D(G(z))]$$

$$= \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_z p_z(z) \log(1 - D(G(z))) dz$$

$$x = G(z) \Rightarrow z = G^{-1}(x) \Rightarrow dz = (G^{-1})'(x) dx$$

$$\Rightarrow p_g(x) = p_z(G^{-1}(x))(G^{-1})'(x)$$

$$= \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_x p_z(G^{-1}(x)) \log(1 - D(x))(G^{-1})'(x) dx$$

$$= \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_x p_g(x) \log(1 - D(x)) dx$$
\[ V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_{z}(z)}[1 - \log D(G(z))] \]

\[ = \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_z p_{z}(z) \log(1 - D(G(z))) dz \]

\[ x = G(z) \Rightarrow z = G^{-1}(x) \Rightarrow dz = (G^{-1})'(x) dx \]

\[ \Rightarrow p_{g}(x) = p_{z}(G^{-1}(x))(G^{-1})'(x) \]

\[ = \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_x p_{z}(G^{-1}(x)) \log(1 - D(x))(G^{-1})'(x) dx \]

\[ = \int_x p_{\text{data}}(x) \log(D(x)) dx + \int_x p_g(x) \log(1 - D(x)) dx \]

\[ = \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) dx \]
Understanding the objective function

\[ \max_D V(D, G) = \max_D \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) \, dx \]
Understanding the objective function

\[
\max_{D} V(D, G) = \max_{D} \int_{x} p_{data}(x) \log(D(x)) + p_{g}(x) \log(1 - D(x)) \, dx
\]

\[
\frac{\partial}{\partial D(x)}(p_{data}(x) \log(D(x)) + p_{g}(x) \log(1 - D(x))) = 0
\]
Understanding the objective function

$$\max_D V(D, G) = \max_D \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) \, dx$$

$$\frac{\partial}{\partial D(x)} (p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x))) = 0$$

$$\Rightarrow \frac{p_{\text{data}}(x)}{D(x)} - \frac{p_g(x)}{1 - D(x)} = 0$$
Understanding the objective function

\[ \max_D V(D, G) = \max_D \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) \, dx \]

\[ \frac{\partial}{\partial D(x)} (p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x))) = 0 \]

\[ \Rightarrow \frac{p_{\text{data}}(x)}{D(x)} - \frac{p_g(x)}{1 - D(x)} = 0 \]

\[ \Rightarrow D(x) = \frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)} \]
Suppose the discriminator is optimal $D^*_G(x)$, the optimal generator makes: $p_{data}(x) = p_g(x)$

$$\Rightarrow D^*_G(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)}$$
Understanding the objective function

\[ C(G) = \max_D V(G, D) \]
Understanding the objective function

\[ C(G) = \max_D V(G, D) \]

\[ = \max_D \int_x p_{data}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) \, dx \]
Understanding the objective function

\[ C(G) = \max_D V(G, D) \]

\[ = \max_D \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) \, dx \]

\[ = \int_x p_{\text{data}}(x) \log(D_G^*(x)) + p_g(x) \log(1 - D_G^*(x)) \, dx \]
Understanding the objective function

\[ C(G) = \max_D V(G, D) \]

\[ = \max_D \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) dx \]

\[ = \int_x p_{\text{data}}(x) \log(D^*_G(x)) + p_g(x) \log(1 - D^*_G(x)) dx \]

\[ = \int_x p_{\text{data}}(x) \log\left(\frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)}\right) + p_g(x) \log\left(\frac{p_g(x)}{p_{\text{data}}(x) + p_g(x)}\right) dx \]
Understanding the objective function

\[ C(G) = \max_D V(G, D) \]

\[ = \max_D \int_x p_{data}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) dx \]

\[ = \int_x p_{data}(x) \log(D_G^*(x)) + p_g(x) \log(1 - D_G^*(x)) dx \]

\[ = \int_x p_{data}(x) \log\left(\frac{p_{data}(x)}{p_{data}(x) + p_g(x)}\right) + p_g(x) \log\left(\frac{p_g(x)}{p_{data}(x) + p_g(x)}\right) dx \]

\[ = \int_x p_{data}(x) \log\left(\frac{p_{data}(x)}{2}\right) + p_g(x) \log\left(\frac{p_g(x)}{2}\right) dx - \log(4) \]
Understanding the objective function

\[ C(G) = \max_D V(G, D) \]

\[ = \max_D \int_x p_{\text{data}}(x) \log(D(x)) + p_g(x) \log(1 - D(x)) \, dx \]

\[ = \int_x p_{\text{data}}(x) \log(D_G^*(x)) + p_g(x) \log(1 - D_G^*(x)) \, dx \]

\[ = \int_x p_{\text{data}}(x) \log\left(\frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)}\right) + p_g(x) \log\left(\frac{p_g(x)}{p_{\text{data}}(x) + p_g(x)}\right) \, dx \]

\[ = \int_x p_{\text{data}}(x) \log\left(\frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)}\right) + p_g(x) \log\left(\frac{p_g(x)}{p_{\text{data}}(x) + p_g(x)}\right) \, dx - \log(4) \]

\[ = KL[p_{\text{data}}(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] + KL[p_g(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] - \log(4) \]
Understanding the objective function

\[ C(G) = KL[p_{\text{data}}(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] + KL[p_g(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] - \log(4) \geq 0 \]
Understanding the objective function

\[ C(G) = KL[p_{data}(x)\|\frac{p_{data}(x)+p_{g}(x)}{2}] + KL[p_{g}(x)\|\frac{p_{data}(x)+p_{g}(x)}{2}] \geq 0 \]

\[ + KL[p_{data}(x)\|\frac{p_{data}(x)+p_{g}(x)}{2}] \geq 0 \]

\[ \min_{G} C(G) = 0 + 0 - \log(4) = -\log(4) \]
Understanding the objective function

\[ C(G) = KL[p_{data}(x) \| \frac{p_{data}(x) + p_{g}(x)}{2}] + KL[p_{g}(x) \| \frac{p_{data}(x) + p_{g}(x)}{2}] \geq 0 \]

\[ \min_{G} C(G) = 0 + 0 - \log(4) = -\log(4) \]

\[ KL[p_{data}(x) \| \frac{p_{data}(x) + p_{g}(x)}{2}] = 0 \]
Understanding the objective function

\[ C(G) = KL[p_{\text{data}}(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] + KL[p_g(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] - \log(4) \geq 0 \geq 0 \]

\[ \min_G C(G) = 0 + 0 - \log(4) = -\log(4) \]

\[ KL[p_{\text{data}}(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] = 0 \]

when \[ p_{\text{data}}(x) = \frac{p_{\text{data}}(x) + p_g(x)}{2} \]
Understanding the objective function

\[ C(G) = KL[p_{\text{data}}(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] + KL[p_g(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] - \log(4) \geq 0 \geq 0 \]

\[ \min_G C(G) = 0 + 0 - \log(4) = -\log(4) \]

\[ KL[p_{\text{data}}(x) \| \frac{p_{\text{data}}(x) + p_g(x)}{2}] = 0 \]

when \( p_{\text{data}}(x) = \frac{p_{\text{data}}(x) + p_g(x)}{2} \)

\[ \Rightarrow p_{\text{data}}(x) = p_g(x) \]
KL (Kullback-Leibler) divergence

- Jensen-Shannon Divergence (symmetric KL):

\[ \text{JSD}(P \| Q) = \frac{1}{2} D_{KL}(P \| M) + \frac{1}{2} D_{KL}(Q \| M), \]

\[ M = \frac{1}{2} (P + Q) \]
Generator $G$, Discriminator $D$

$$V = \mathbb{E}_{x \sim P_{data}}[\log D(x)] + \mathbb{E}_{x \sim P_{G}}[\log(1 - D(x))]$$
Summary:

- Generator $G$, Discriminator $D$
- Looking for $G^*$ such that

$$V = \mathbb{E}_{x \sim P_{\text{data}}} [\log D(x)] + \mathbb{E}_{x \sim P_{G}} [\log (1 - D(x))]$$

$$G^* = \arg \min_G \max_D V(G, D)$$
Summary:

- Generator $G$, Discriminator $D$
- Looking for $G^*$ such that

$$G^* = \arg \min_G \max_D V(G, D)$$

- Given $G$, $\max_D V(G, D)$

$$= -2\log(2) + 2\text{JSD}(P_{\text{data}}(x) \| P_G(x))$$
Generator $G$, Discriminator $D$

Looking for $G^*$ such that

$$G^* = \arg \min_G \max_D V(G, D)$$

Given $G$, $\max_D V(G, D)$

$$= -2\log(2) + 2\text{JSD}(P_{data}(x) \| P_G(x))$$

What is the optimal $G$? It is $G$ that makes $\text{JSD}$ smallest $= 0$:

$$P_G(x) = P_{data}(x)$$

$$V = \mathbb{E}_{x \sim P_{data}}[\log D(x)] + \mathbb{E}_{x \sim P_G}[\log (1 - D(x))]$$
<table>
<thead>
<tr>
<th>Caption</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>a pitcher is about to throw the ball to the batter</td>
<td></td>
</tr>
<tr>
<td>a group of people on skis stand in the snow</td>
<td></td>
</tr>
<tr>
<td>a man in a wet suit riding a surfboard on a wave</td>
<td></td>
</tr>
</tbody>
</table>
## Text to Image - Results

From CY Lee lecture

<table>
<thead>
<tr>
<th>Caption</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>this flower has white petals and a yellow stamen</td>
<td><img src="image1.png" alt="Images of white flowers with a yellow stamen" /></td>
</tr>
<tr>
<td>the center is yellow surrounded by wavy dark purple petals</td>
<td><img src="image2.png" alt="Images of purple flowers with wavy petals" /></td>
</tr>
<tr>
<td>this flower has lots of small round pink petals</td>
<td><img src="image3.png" alt="Images of pink flowers with small round petals" /></td>
</tr>
</tbody>
</table>

Project topic: Code and data are all on web, many possibilities!
<table>
<thead>
<tr>
<th>Caption</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>this flower has white petals and a yellow stamen</td>
<td><img src="image1" alt="Images of flowers with white petals and a yellow stamen" /></td>
</tr>
<tr>
<td>the center is yellow surrounded by wavy dark purple petals</td>
<td><img src="image2" alt="Images of flowers with yellow center and wavy dark purple petals" /></td>
</tr>
<tr>
<td>this flower has lots of small round pink petals</td>
<td><img src="image3" alt="Images of flowers with lots of small round pink petals" /></td>
</tr>
</tbody>
</table>

Project topic: Code and data are all on web, many possibilities!

"red flower with black center"

From CY Lee lecture
"red flower with black center"

From CY Lee lecture

<table>
<thead>
<tr>
<th>Caption</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>this flower has white petals and a yellow stamen</td>
<td>![Images]</td>
</tr>
<tr>
<td>the center is yellow surrounded by wavy dark purple petals</td>
<td>![Images]</td>
</tr>
<tr>
<td>this flower has lots of small round pink petals</td>
<td>![Images]</td>
</tr>
</tbody>
</table>

Project topic: Code and data are all on web, many possibilities!
Real images (CIFAR-10)  Generated images
Source Code

• Original paper (theano):
  – https://github.com/goodfeli/adversarial
• Tensorflow implementation:
  – https://github.com/ckmarkoh/GAN-tensorflow
In practice...

- Given $G$, how to compute $\max_D V(G, D)$?
  - Sample $\{x^1, \ldots, x^m\}$ from $P_{data}$
  - Sample $\{x^*_1, \ldots, x^*_m\}$ from generator $P_G$

Maximize:

$$V' = \frac{1}{m \sum_{i=1}^{m} \log D(x^i)} + \frac{1}{m \sum_{i=1}^{m} \log (1 - D(x^*_i))}$$

$$V = \mathbb{E}_{x \sim P_{data}}[\log D(x)] + \mathbb{E}_{x \sim P_G}[\log (1 - D(x))]$$

Credit: Mark Chang
In practice ...

- Given $G$, how to compute $\max_D V(G, D)$?
  - Sample $\{x^1, \ldots, x^m\}$ from $P_{data}$
  - Sample $\{x^*1, \ldots, x^*m\}$ from generator $P_G$

Maximize:

$$V' = \frac{1}{m \sum_{i=1}^m \log D(x^i)} + \frac{1}{m \sum_{i=1}^m \log (1-D(x^i))}$$

This is what a Binary Classifier do

Output is $D(x)$ Minimize Cross-entropy

If $x$ is a positive example  ➔ Minimize $-\log D(x)$
If $x$ is a negative example ➔ Minimize $-\log(1-D(x))$
In practice...

- Given $G$, how to compute $\max_D V(G, D)$?
  - Sample $\{x^1, \ldots, x^m\}$ from $P_{\text{data}}$
  - Sample $\{x^{*1}, \ldots, x^{*m}\}$ from generator $P_G$

Maximize:

$$V' = \frac{1}{m \sum_{i=1}^{m} \log D(x^i)} + \frac{1}{m \sum_{i=1}^{m} \log (1 - D(x^{*i}))}$$

This is what a Binary Classifier do

Output is $D(x)$ Minimize Cross-entropy

- If $x$ is a positive example $\Rightarrow$ Minimize $-\log D(x)$
- If $x$ is a negative example $\Rightarrow$ Minimize $-\log (1 - D(x))$
In practice ...

Given $G$, how to compute $\max_D V(G, D)$?
- Sample $\{x^1, x^m\}$ from $P_{data}$
- Sample $\{x^{'1}, x^{'m}\}$ from generator $P_G$

Maximize:

$$V' = \frac{1}{m \sum_{i=1}^{m} \log D(x^i)} + \frac{1}{m \sum_{i=1}^{m} \log(1 - D(x^{'i}))}$$

This is what a Binary Classifier do

Output is $D(x)$ Minimize Cross-entropy
- If $x$ is a positive example Minimize $-\log D(x)$
- If $x$ is a negative example Minimize $-\log(1 - D(x))$

$$V = \mathbb{E}_{x \sim P_{data}}[\log D(x)] + \mathbb{E}_{x \sim P_G}[\log(1 - D(x))]$$